On Wielandt's Inequality and its Application to the Asymptotic Distribution of the Eigenvalues of a Random Symmetric Matrix

by

Morris L. Eaton, University of Minnesota
and
David E. Tyler, Rutgers University

School of Statistics, University of Minnesota
Technical Report No. 518
September 1988

On Wielandt's Inequality and its Application to the Asymptotic Distribution of the Eigenvalues of a Random Symmetric Matrix

by

Morris L. Eaton[1]
School of Statistics
University of Minnesota
Minneapolis, Minnesota 55455

and

David E. Tyler[2]
Department of Statistics
Rutgers University
New Brunswick, New Jersey 08903

[1]Research supported by NSF grant DMS 88–00284.

[2]Research supported by NSF grant DMS 87–01198.

SUMMARY

A relatively obscure eigenvalue inequality due to Wielandt is used to give a simple derivation of the asymptotic distribution of the eigenvalues of a random symmetric matrix. The asymptotic distributions are obtained under a fairly general setting. An application of the general theory to the bootstrap distribution of the eigenvalues of the sample covariance matrix is given.

AMS 1980 subject classifications. Primary 62H25, 62E20.

Key words and phrases. Bootstrap, covariance matrix, eigenvalues, random symmetric matrices.

1. Introduction and Summary.

The derivation of the asymptotic distribution of the eigenvalues of a random symmetric matrix arises in many papers in multivariate analysis. Although the main idea behind most of the derivations is quite basic, i.e., the expansion of the "sample" roots about the "population" roots, the derivations themselves are often quite involved and difficult to follow. These complications are primarily due to the mathematical rather than statistical nature of the eigenvalue problem.

One of the main objectives of this paper is to introduce a simple method for obtaining the asymptotic distribution of the eigenvalue of random symmetric matrices. The method is based upon a relatively obscure eigenvalue inequality due to Wielandt (1967).

Obtaining the asymptotic distribution of eigenvalues by expanding the "sample" roots about the "population" roots becomes even more cumbersome when the "population" roots vary, e.g., see Tyler (1983). This case arises when considering local alternatives to hypotheses on "population" covariance matrices, and it also arises when considering the bootstrap distribution of eigenvalues, see Beran and Srivastava (1985). For this case, the use of Wielandt's eigenvalue inequality again provides a fairly simple method for obtaining the asymptotic distribution of the roots.

This paper is organized as follows. Wielandt's eigenvalue inequality is stated and discussed in section 2. General results on the asymptotic distribution of eigenvalues of random symmetric matrices are presented in section 3. The case when the "population" roots vary is treated in section 4. Finally, the results of sections 3 and 4 are applied in section 5 to obtain results on the asymptotic behavior of the bootstrap distribution of the eigenvalues of the sample covariance matrix. The results on the bootstrap extend the work of Beran and Srivastava (1985, 1987).

2. Wielandt's Eigenvalue Inequality.

Consider a symmetric matrix

$$A = \begin{bmatrix} B & C \\ C' & D \end{bmatrix} \tag{2.1}$$

where A is $p \times p$, B is $q \times q$ and D is $r \times r$. Let $\rho^2(C)$ denote the largest eigenvalue of CC', and let $\alpha_1 \geq \dots \geq \alpha_p$, $\beta_1 \geq \dots \geq \beta_q$ and $\delta_1 \geq \dots \geq \delta_r$ be the ordered eigenvalues of A, B and D, respectively.

THEOREM 2.1 (WIELANDT). If $\beta_q > \delta_1$, then

$$0 \leq \alpha_j - \beta_j \leq \rho^2(C)/(\beta_j - \delta_1), \quad j = 1,\dots,q \tag{2.2}$$

and

$$0 \leq \delta_{r-i} - \alpha_{p-i} \leq \rho^2(C)/(\beta_q - \delta_{r-i}), \ i = 0,\dots,r-1. \tag{2.3}$$

The first set of inequalities (2.2) is given in Wielandt's (1967) lecture notes on page 120, but only when A is positive definite. The inequalities follow immediately for any symmetric A by replacing A by $A+\delta I$ where $\delta > -\delta_r$ and noting that $A+\delta I$ is positive definite and the δ term cancels in (2.2). The second set of inequalities (2.3) follow from the first by multiplying A by -1.

The first inequality in (2.2) is simply a partial restatement of result (1f.2.13) in Rao (1973) which he refers to as a Sturmian Separation Theorem, and which is referred to by Wielandt (1967) on page 117 as the Interlacement Theorem. The second inequality in (2.2) is apparently a novel result of Wielandt. Two interesting features of this inequality deserve to be noted. First, note that the matrix A can be viewed as a perturbation of a block diagonal matrix, namely $A = A_0 + E$ where

$$A_0 = \begin{bmatrix} B & 0 \\ 0 & D \end{bmatrix} \quad \text{and} \quad E = \begin{bmatrix} 0 & C \\ C' & 0 \end{bmatrix}. \tag{2.4}$$

By Wielandt's inequality, the eigenvalues of A_0 are perturbed quadratically in E when A_0 is perturbed linearly in E. It is well known that in general, eigenvalues are only perturbed linearly when the matrix is perturbed linearly. The quadratic perturbation obtained in Wielandt's inequality is due to the special structure of E relative to A_0. This quadratic perturbation result can also be obtained in a somewhat more cumbersome way by using perturbation techniques, for example as described in Chapter 2 of Kato (1976), and by observing that due to the special structure of E relative to A_0, the linear term is zero.

The other interesting feature of Wielandt's inequality is that it not only shows that the perturbation of the eigenvalues are of quadratic order, but it also gives a bound which shows how the perturbation is related to the separation of the eigenvalues of B and D. Most perturbation techniques, such as Taylor series expansions, give only an order of perturbation in E rather than bounds. Bounds of quadratic order on the perturbed eigenvalues can be obtained by using perturbation techniques described in Chapter 2 of Kato (1976), as is done in section 6 of Tyler (1983). In light of Wielandt's inequality, this approach is unnecessary, especially since it is more complicated and gives weaker bounds.

Because of its central role in this paper and its relative unavailability in the literature, a brief but complete proof of Wielandt's inequality is given below. The proof relies heavily on the following lemma which is given as problem 1.9 on page 68 in Rao (1973). Wielandt (1967) refers to this lemma as Weyl's Theorem (page 114), and refers to the corollary stated after the lemma as Aronszajn's Theorem (page 119). The simple proof of the corollary is due to Wielandt.

LEMMA 2.1 (WEYL). Let $T = R+S$ where R and S are $k\times k$ symmetric matrices, and let $t_1 \geq \dots \geq t_k$, $r_1 \geq \dots \geq r_k$, and $s_1 \geq \dots \geq s_k$ be the ordered eigenvalues of T, R and S, respectively, then

$$t_j \leq \begin{cases} r_j+s_1 \\ \vdots \\ r_1+s_j \end{cases} \quad \text{and} \quad t_j \geq \begin{cases} r_j+s_k \\ \vdots \\ r_k+s_j \end{cases} . \tag{2.5}$$

COROLLARY 2.1. Suppose in (2.1) $A \geq 0$, i.e., A is positive semi-definite, then

$$\beta_i \leq \alpha_i \leq \begin{cases} \beta_i+\delta_1 \\ \vdots \\ \beta_1+\delta_i \end{cases}$$

for $i = 1,\dots,p$, where $\beta_i = 0$ for $i > q$ and $\delta_i = 0$ for $i > r$.

Proof of Corollary. Since A is symmetric positive semi-definite, it has a symmetric positive definite square root $A^{1/2}$. Let $A^{1/2} = [F \;\; G]$ where F is $p\times q$ and G is $p\times r$, and so A can be expressed as

$$A = \begin{bmatrix} F'F & F'G \\ G'F & G'G \end{bmatrix} \quad \text{and} \quad A = FF' + GG'.$$

The corollary follows from the lemma by noting that FF' and $F'F = B$ have the same eigenvalues apart from zeros. □

Proof of Theorem 2.1. As noted previously, it only remains to show the second inequality in (2.2) holds. Since the result is invariant under the transformation $A \to A+\gamma I$ and under the transformation $B \to P'BP$, $D \to Q'DQ$ and $C \to P'CQ$ for orthogonal matrices P and Q, it can be assumed w.l.o.g. that $B = \mathrm{diag}\{\beta_1,\ldots,\beta_q\}$, $D = \mathrm{diag}\{\delta_1,\ldots,\delta_q\}$ and $\delta_1 = 0$. Note that the first inequality in (2.2) then implies $\alpha_i \geq \beta_i > 0$ for $i = 1,\ldots,q$.

Let B be the $q\times q$ matrix $\tilde{B} = \mathrm{diag}\{\beta_1,\ldots,\beta_j,\ldots,\beta_j\}$ for a fixed j. Also, let

$$\tilde{A} = \begin{pmatrix} \tilde{B} & C \\ C' & 0 \end{pmatrix} \quad \text{and hence} \quad \tilde{A}^2 = \begin{pmatrix} \tilde{B}^2+CC' & \tilde{B}C \\ C'\tilde{B} & C'C \end{pmatrix}.$$

Let $\tilde{\alpha}_1 \geq \ldots \geq \tilde{\alpha}_p$ be the ordered eigenvalues of $\tilde{A}$. Now since $D \leq 0$ and $\tilde{B} \geq B$, it follows that $\tilde{A} \geq A$ and hence $\tilde{\alpha}_i \geq \alpha_i$, $i = 1,\ldots,p$. (The notation $M_1 \geq M_2$ means M_1-M_2 is positive semi–definite.) Let $\pi_1 \geq \ldots \geq \pi_p$ be the ordered eigenvalues of $\tilde{A}^2$ and so they represent the ordered values of $\tilde{\alpha}_i^2$, $i = 1,\ldots,p$. Note that π_i is not necessarily equal to $\tilde{\alpha}_i^2$ since $\tilde{A}$ is not positive semi–definite. However, since $\tilde{\alpha}_i \geq \alpha_i > 0$ for $i = 1,\ldots,q$, it follows that $\pi_i \geq \tilde{\alpha}_i^2 \geq \alpha_i^2$ for $i = 1,\ldots,q$, and in particular $\pi_j \geq \alpha_j^2$.

An upper bound on π_j can be obtained by first applying Corollary 2.1 to $\tilde{A}^2$. This gives $\pi_j \leq \tilde{\beta}_j^2 + \rho^2(C)$, where $\tilde{\beta}_j^2$ is the jth largest root of $\tilde{B}^2+CC'$. Application of Lemma 2.1 to $\tilde{B}^2+CC'$ then gives $\tilde{\beta}_j^2 \leq \beta_j^2 + \rho^2(C)$ and so $\pi_j \leq \beta_j^2 + 2\rho^2(C)$. Putting the two inequalities for π_j together yields $\alpha_j^2 \leq \beta_j^2 + 2\rho^2(C)$ or $(\alpha_j-\beta_j)(\alpha_j+\beta_j) = \alpha_j^2-\beta_j^2 \leq 2\rho^2(C)$. However, since $\alpha_j \geq \beta_j$, the desired result $\alpha_j-\beta_j \leq \rho^2(C)/\beta_j$ follows. □

3. Asymptotic Results for Eigenvalues of Random Symmetric Matrices.

Consider a sequence of random matrices S_n, $n = 1,2,\ldots$, in $\mathcal{S}_p$, the set of $p \times p$ real symmetric matrices, and assume that

$$W_n = n^{1/2}(S_n - \Sigma) \to_d W \tag{3.1}$$

with $\Sigma \in \mathcal{S}_p$ and hence $W \in \mathcal{S}_p$. Given $M \in \mathcal{S}_p$, let $\varphi(M) = (\varphi_1(M),\ldots,\varphi_p(M))$ be the vector of ordered eigenvalues of M. The dependence of φ on the dimension parameter is suppressed and the same symbol φ is used for the vector of ordered eigenvalues of symmetric matrices of different dimensions. In this section, the asymptotic distribution of

$$X_n = n^{1/2}\{\varphi(S_n) - \varphi(\Sigma)\} \in \mathbb{R}^p \tag{3.2}$$

is studied. Without loss of generalizity, Σ is taken to be diagonal, in particular $\mathrm{diag}\{\varphi(\Sigma)\}$. In what follows, the choice of norms in $\mathbb{R}^p$ and in $\mathcal{S}_p$ is irrelevant. The notation $T_n = O_p(b_n^{-1})$ implies for any sequence of positive numbers a_n with $a_n \to 0$, $a_n b_n \|T_n\| \to 0$ in probability. The notation $T_n = o_p(b_n^{-1})$ implies $b_n\|T_n\| \to 0$ in probability.

3.1. A Basic Lemma.

Partition Σ and S_n as

$$\Sigma = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix} \quad \text{and} \quad S_n = \begin{pmatrix} T_n & U_n \\ U_n' & V_n \end{pmatrix} \tag{3.3}$$

where Σ_{11} is the $q \times q$ diagonal matrix $\Sigma_{11} = \mathrm{diag}\{\varphi_1(\Sigma),\ldots,\varphi_q(\Sigma)\}$, Σ_{22} is the $r \times r$ diagonal matrix $\Sigma_{22} = \mathrm{diag}\{\varphi_{q+1}(\Sigma),\ldots,\varphi_p(\Sigma)\}$ and $p = q+r$. The matrices T_n, V_n and U_n are $q \times q$, $r \times r$ and $r \times q$, respectively.

LEMMA 3.1. If $\varphi_q(\Sigma) > \varphi_{q+1}(\Sigma)$, then

$$Y_n = \varphi(S_n) - \begin{bmatrix}\varphi(T_n)\\ \varphi(V_n)\end{bmatrix} \text{ is } O_p(n^{-1}).$$

Proof. Let $A_n = \{\varphi_q(T_n) > \varphi_1(V_n)\}$. Since φ is a continuous function and from (3.1), $T_n \to_p \Sigma_{11}$ and $V_n \to_p \Sigma_{22}$, it follows that $\varphi_q(T_n) \to_p \varphi_q(\Sigma_{11}) = \varphi_q(\Sigma)$ and $\varphi_1(V_n) \to_p \varphi_1(\Sigma_{22}) = \varphi_{q+1}(\Sigma)$. Thus, Prob $(A_n) \to 1$, and so attention can be restricted to A_n, $n = 1,2,\ldots$. For $S_n \in A_n$, Wielandt's Theorem (Theorem 2.1) implies

$$|\varphi_i(S_n) - \varphi_i(T_n)| < \rho^2(U_n)/\{\varphi_q(T_n) - \varphi_1(V_n)\}. \tag{3.4}$$

Now, by (3.1), $U_n = O_p(n^{-1/2})$ and since ρ is continuous it follows that $\rho^2(U_n) = O_p(n^{-1})$. The top part of the lemma then follows from (3.4) since $\varphi_q(T_n) - \varphi_1(V_n) \to_p \varphi_q(\Sigma) - \varphi_{q+1}(\Sigma) > 0$. The proof of the bottom part is analogous to the top. □

3.2. The Main Theorems.

Let $d_1 > d_2 > \ldots > d_k$ represent the distinct eigenvalues of Σ with the multiplicity of d_i being p_i, $i = 1,\ldots,k$, and hence $p_1+\ldots+p_k = p$. Let I_i be the $p_i \times p_i$ identity matrix and partition Σ and S_n as

$$\Sigma = \begin{bmatrix} d_1 I_1 & 0 & \cdots & 0 \\ 0 & d_2 I_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & d_k I_k \end{bmatrix} \text{ and } S_n = \begin{bmatrix} S_{n,11} & S_{n,11} & \cdots & S_{n,1k} \\ S_{n,21} & S_{n,22} & \cdots & S_{n,2k} \\ \vdots & \vdots & & \vdots \\ S_{n,k1} & S_{n,k2} & \cdots & S_{n,kk} \end{bmatrix} \tag{3.5}$$

where $S_{n,ij}$ is $p_i \times p_j$, $i,j = 1,\ldots,k$. By applying Lemma 3.1 $k-1$ times, the following asymptotic equivalence result is obtained. The vector $e_i \in \mathbb{R}^{p_i}$ is the vector of ones, $i = 1,\ldots,k$.

THEOREM 3.1. In the notation above, $n^{1/2}\{\varphi(S_n) - \varphi(\Sigma)\} = Z_n + R_n$, where

$$Z_n = n^{1/2}\begin{bmatrix} \varphi(S_{n,11})-d_1e_1 \\ \vdots \\ \varphi(S_{n,kk})-d_ke_k \end{bmatrix}$$

and the remainder term R_n is $O_p(n^{-1/2})$.

The asymptotic distribution of the leading term Z_n can be readily obtained from (3.1). Analogous to the partitioning of S_n, let $W = \{W_{ij}\}$ represent the partitioning of W in blocks of order $p_i \times p_j$, and hence

$$\tilde{W}_n = n^{1/2}\begin{bmatrix} S_{n,11}-d_1I_1 \\ \vdots \\ S_{n,kk}-d_kI_k \end{bmatrix} \to_d \tilde{W} = \begin{bmatrix} W_{11} \\ \vdots \\ W_{kk} \end{bmatrix}. \tag{3.6}$$

Now, on the space $\mathcal{S}_{p_1} \oplus \mathcal{S}_{p_2} \oplus \ldots \oplus \mathcal{S}_{p_k}$, where $\tilde{W}$ takes its values, the function

$$H(\tilde{W}) = \begin{bmatrix} \varphi(W_{11}) \\ \vdots \\ \varphi(W_{kk}) \end{bmatrix} \tag{3.7}$$

is continuous, and hence $H(\tilde{W}_n) \to_d H(W)$. However, since $\varphi\{n^{1/2}(S_{n,ii} - d_i I_i)\} = n^{1/2}\{\varphi(S_{n,ii}) - d_i e_i\}$, $i = 1,\ldots,k$, the following theorem is obtained.

THEOREM 3.2. In the above notation, $\sqrt{n}\{\varphi(S_n) - \varphi(\Sigma)\} = H(\tilde{W}_n) + R_n$, where R_n is $O_p(n^{-1/2})$ and $H(\tilde{W}_n) \to_d H(\tilde{W})$.

Thus, the asymptotic distribution of the roots of S_n is found by calculating the distribution of $H(\tilde{W})$.

4. Asymptotic Behavior under a More General Setting.

Consider now two sequences of random matrices S_n and Σ_n, $n = 1,2,\ldots,$ both in $\mathcal{S}_p$, and assume that

$$W_n = n^{1/2}(S_n - \Sigma_n) = O_p(1). \tag{4.1}$$

As a special case, Σ_n may be a sequence of nonrandom matrices, and in particular if Σ_n does not depend on n and W_n converges in distribution, then this reduces to the setting in section 3. In this section, the asymptotic behavior of

$$X_n = n^{1/2}\{\varphi(S_n) - \varphi(\Sigma_n)\} \in \mathbb{R}^p \tag{4.2}$$

is studied. Using the spectral value decomposition, express $\Sigma_n = P_n'\Delta_n P_n$ where P_n is an orthogonal matrix and Δ_n is the diagonal matrix with diagonal entries $\varphi_1(\Sigma_n),\ldots,\varphi_p(\Sigma_n)$ respectively. Define $S_n^0 = P_n S_n P_n'$ and note that $X_n = n^{1/2}\{\varphi(S_n^0) - \varphi(\Delta_n)\}$.

4.1. A Basic Lemma.

Partition Δ_n and S_n^0 as

$$\Delta_n = \begin{pmatrix} \Delta_{11,n} & 0 \\ 0 & \Delta_{22,n} \end{pmatrix} \quad \text{and} \quad S_n^0 = \begin{pmatrix} T_n & U_n \\ U_n' & V_n \end{pmatrix} \tag{4.3}$$

where the dimensions are the same as in (3.3).

LEMMA 4.1. If $a_n\{\varphi_q(\Sigma_n) - \varphi_{q+1}(\Sigma_n)\} \to_p \infty$ for some increasing sequence of positive numbers $a_n \to \infty$ with $a_n^{-1} = O_p(n^{-1/2})$, then

$$Y_n = \varphi(S_n) - \begin{bmatrix} \varphi(T_n) \\ \varphi(V_n) \end{bmatrix} \text{ is } o_p(a_n/n).$$

Proof. Let $A_n = \{\varphi_q(T_n) > \varphi_1(V_n)\}$ and $B_n = \{\varphi_q(\Delta_{11,n}) > \varphi_1(\Delta_{22,n})\}$. By the condition in the lemma, it readily follows that Prob $(B_n) \to 1$. Condition (4.1) implies $P_nW_nP_n' = O_p(1)$ and hence $n^{1/2}(T_n - \Delta_{11,n}) = O_p(1)$ and $n^{1/2}(V_n - \Delta_{22,n}) = O_p(1)$. Application of Lemma 2.1 thus gives $\varphi_q(T_n) = \varphi_q(\Delta_{11,n}) + O_p(n^{-1/2})$ and $\varphi_1(V_n) = \varphi_1(\Delta_{22,n}) + O_p(n^{-1/2})$, which implies $a_n\{\varphi_q(T_n) - \varphi_1(V_n)\} \to_p \infty$, and so Prob $(A_n) \to 1$.

Attention is now restricted to A_n, $n = 1,2,\ldots$. For $S_n \in A_n$, Wielandt's Theorem and the identity $\varphi(S_n) = \varphi(S_n^*)$ gives

$$|\varphi_i(S_n) - \varphi_i(T_n)| < \rho^2(U_n)/\{\varphi_q(T_n) - \varphi_1(V_n)\}. \tag{4.4}$$

The numerator is $O_p(n^{-1})$ since $P_nW_nP_n'$ is $O_p(n^{-1/2})$ and so U_n is $O_p(n^{-1/2})$. It already has been shown that $a_n\{\varphi_q(T_n) - \varphi_1(V_n)\} \to_p \infty$ and hence the right–hand side of (4.4) is $o_p(a_n/n)$. The proof of the bottom part of the theorem is analogous. □

4.2. The Main Theorems.

Partition the matrices Δ_n and S_n^0, respectively, as

$$\Delta_n = \begin{bmatrix} \Delta_{n,1} & 0 & \cdots & 0 \\ 0 & \Delta_{n,2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Delta_{n,k} \end{bmatrix} \text{ and } S_n^0 = \begin{bmatrix} S_{n,11}^0 & S_{n,12}^0 & \cdots & S_{n,1k}^0 \\ S_{n,21}^0 & S_{n,22}^0 & \cdots & S_{n,2k}^0 \\ \vdots & \vdots & \ddots & \vdots \\ S_{n,k1}^0 & S_{n,k2}^0 & \cdots & S_{n,kk}^0 \end{bmatrix}, \tag{4.5}$$

where the dimensions are analogous to those in (3.5). For example, $\Delta_{n,i}$ and $S^0_{n,ii}$ are $p_i \times p_i$. Application of Lemma 4.1 k–1 times with $a_n = n^{1/2}$ gives the following result.

THEOREM 4.1. If $n^{1/2}\{\varphi_{p_i}(\Delta_{n,i}) - \varphi_1(\Delta_{n,i+1})\} \to_p \infty$ for $i = 1,2,\ldots,k-1$, then $X_n = n^{1/2}\{\varphi(S_n) - \varphi(\Sigma_n)\} = Z_n + R_n$, where

$$Z_n = n^{1/2}\begin{bmatrix} \varphi(S^0_{n,11}) - \varphi(\Delta_{n,1}) \\ \vdots \\ \varphi(S^0_{n,kk}) - \varphi(\Delta_{n,k}) \end{bmatrix}$$

and the remainder term R_n is $o_p(1)$.

Thus, the asymptotic distribution of X_n, if it exists, is the same as the asymptotic distribution of Z_n. Even if the asymptotic distribution does not exist, Z_n represents a simpler asymptotically equivalent variate. The term Z_n can be re–expressed as follows. Let $d_{n,i}$ represent the average of the p_i eigenvalues in $\Delta_{n,i}$, define $D_{n,i} = n^{1/2}(\Delta_{n,i} - d_{n,i}I_i)$. Also, define $W^0_n = P_n W_n P_n'$ with $W^0_n = \{W^0_{n,ij}\}$ representing the partitioning of W^0_n in blocks of order $p_i \times p_j$, and let

$$\tilde{W}^0_n = n^{1/2}\begin{bmatrix} S^0_{n,11} - \Delta_{n,1} \\ \vdots \\ S^0_{n,kk} - \Delta_{n,k} \end{bmatrix} = \begin{bmatrix} W^0_{n,11} \\ \vdots \\ W^0_{n,kk} \end{bmatrix} \text{ and } \tilde{D}_n = \begin{bmatrix} D_{n,1} \\ \vdots \\ D_{n,k} \end{bmatrix}. \tag{4.6}$$

Note that $n^{1/2}\{\varphi(S^0_{n,ii}) - d_{n,i}e_i\} = \varphi\{n^{1/2}(S^0_{n,ii} - d_{n,i}I_i)\} = \varphi(W^0_{n,ii} + D_{n,i})$ and $n^{1/2}\{\varphi(\Delta_{n,i}) - d_{n,i}e_i\} = \varphi(D_{n,i})$, and so using the function H defined in section 3, the following result is obtained.

THEOREM 4.2. In the above notation and under the conditions of Theorem 4.1, $n^{1/2}\{\varphi(S_n) - \varphi(\Sigma_n)\} = \{H(\tilde{W}_n^0 + \tilde{D}_n) - H(\tilde{D}_n)\} + R_n$, where R_n is $o_p(1)$.

For nonrandom Σ_n, Theorem 4.2 can be used to obtain the asymptotic distribution of the roots of S_n under the sequence Σ_n. Suppose $\Sigma_n \to \Sigma$, which without loss of generality is taken as in (3.5). The sequence P_n can be chosen so that $P_n \to I$, and so if

$$W_n = n^{1/2}(S_n - \Sigma_n) \to_d W, \tag{4.7}$$

then $W_n^0 \to_d W$ and hence $\tilde{W}_n^0 \to_d \tilde{W}$. Furthermore if $D_n \to D$, then

$$n^{1/2}\{\varphi(S_n) - \varphi(\Sigma_n)\} \to_d H(\tilde{W} + D) - H(D). \tag{4.8}$$

Note that no condition on the rate at which $\Sigma_n \to \Sigma$ is made in obtaining (4.8). Only conditions on the rates at which the roots of Σ_n approach each other or diverge from each other are needed.

The term $n^{1/2}$ in (3.1) and (4.1) is the most common rate arising in practice. The results of this section and section 3, though, readily generalize if the rate $n^{1/2}$ in (3.1) and (4.2) is replaced by an increasing sequence $c_n \to \infty$. The resulting modification in all the statements and theorems is made by simply replacing n by c_n^2 (except, of course, when n is used as an index or subscript).

Wielandt's Theorem is also valid when A in (2.1) has complex entries and is self–adjoint. Correspondingly, the results of this section and section 3 can be easily extended to the case when S_n, Σ and Σ_n have complex entries and are self–adjoint.

5. An Application to Bootstrapping Eigenvalues.

Let $\{x_i;\ 1 \leq i \leq n\}$ represent a random sample from a distribution with covariance matrix Σ and finite fourth moments. If S_n represents the sample covariance matrix, then

$$W_n = n^{1/2}(S_n - \Sigma) \rightarrow_d W, \tag{5.1}$$

where W has a multivariate normal distribution. Without loss of generality, let $\Sigma = \Delta$ be diagonal and represented as in (3.5). Using the notation established in section 3.2, application of Theorem 3.2 gives

$$X_n = n^{1/2}\{\varphi(S_n) - \varphi(\Sigma)\} \rightarrow_d H(\tilde{W}). \tag{5.2}$$

Let F_n be the sample distribution function of $\{x_i;\ 1 \leq i \leq n\}$. The covariance matrix associated with the distribution F_n is thus S_n. Consider now a random sample $\{x_i^*;\ 1 \leq i \leq n\}$ from the distribution F_n and let S_n^* be the sample covariance matrix of this sample. The idea behind the bootstrap is to use the distribution of $W_n^* = n^{1/2}(S_n^* - S_n)$ under F_n, which is realizable, as a nonparametric estimate of the distribution of $W_n = n^{1/2}(S_n - \Sigma)$. Beran and Srivastava (1985) show that the bootstrap estimate is strongly consistent, that is

$$W_n^* = n^{1/2}(S_n^* - S_n) \rightarrow_{d*} W \quad \text{a.s.} \tag{5.3}$$

The notation $\rightarrow_{d*}$ refers to the weak convergence of the distribution function of W_n^* under F_n to the distribution function of W. Under F_n, S_n is a fixed matrix, and S_n^* is a random matrix. The almost sure statement refers to the underlying product measure on $\{x_i;\ 1 \leq i \leq n\}$.

The nonparametric bootstrap estimate of the distribution of $X_n = n^{1/2}\{\varphi(S_n) - \varphi(\Sigma)\}$ is the distribution of $X_n^* = n^{1/2}\{\varphi(S_n^*) - \varphi(S_n)\}$ under F_n. It is easy to verify that the conditions of Theorem 4.2 are almost surely satisfied and so

$$X_n^* - \{H(\tilde{W}_n^{*0} + \tilde{D}_n) - H(\tilde{D}_n)\} \to_{d^*} 0 \quad \text{a.s.} \tag{5.4}$$

The notation in (5.4) is as follows. Let $S_n = P_n'\Delta_n P_n$ represent the spectral value decomposition of S_n with $\Delta_n = \text{diag}\{\varphi(S_n)\}$. Defined $W_n^{*0} = P_n W_n^* P_n'$ and hence $\tilde{W}_n^{*0}$ is defined accordingly. Partition Δ_n as in (4.5) and then $\tilde{D}_n$ is defined as in (4.6). Since $\varphi_i(A + aI) = \varphi_i(A) + a$, expression (5.4) can be re-expressed as

$$X_n^* - \{H(\tilde{W}_n^{*0} + \tilde{A}_n) - H(\tilde{A}_n)\} \to_{d^*} 0 \quad \text{a.s.}, \tag{5.5}$$

where $A_n = n^{1/2}(\Delta_n - \Delta)$ and $\tilde{A}_n$ is defined accordingly.

Now, since $S_n \to \Sigma = \Delta$ a.s., the sequence P_n can be chosen so that $P_n \to I$ a.s. and hence $\tilde{W}_n^{*0} \to_{d^*} \tilde{W}$ a.s. . The matrices $\tilde{A}_n$ are fixed matrices with respect to F_n, and converge in distribution but not almost surely with respect to the product measure on $\{x_i;\ 1 \le i \le \infty\}$. More specifically,

$$\tilde{A}_n = \begin{pmatrix} A_{n,1} \\ \vdots \\ A_{n,k} \end{pmatrix} \to_d A = \begin{pmatrix} A_1 \\ \vdots \\ A_k \end{pmatrix} \tag{5.6}$$

where $A_{n,i} = n^{1/2}(\Delta_{n,i} - d_i I)$ and so from (5.2) the joint distribution of $A_1, \ldots, A_k$ is the same as the joint distribution of $\text{diag}\{\varphi(W_{11})\}, \ldots, \text{diag}\{\varphi(W_{kk})\}$.

If all the eigenvalues of Σ are distinct, then $H(\tilde{W}_n^{*0} + \tilde{A}_n) - H(\tilde{A}_n) = H(\tilde{W}^{*0})$, and so from (5.5)

$$X_n^* = n^{1/2}\{\varphi(S_n^*) - \varphi(S_n)\} \rightarrow_{d^*} H(\tilde{W}) \quad \text{a.s.} \tag{5.7}$$

Thus, for this case the bootstrap distribution for the eigenvalues are strongly consistent. If only some of the eigenvalues of Σ are distinct, then by the same argument it can be shown that the joint marginal distribution of the bootstap distribution associated with these roots are strongly consistent. However, since $\tilde{A}_n$ does not go to zero almost surely, and does not cancel out in (5.5) when Σ has multiple roots, the marginal bootstrap distribution associated with a multiple root is not consistent. The consistency of the bootstrap for distinct roots was proven by Beran and Srivastava (1985). They also showed the inconsistency of the bootstrap in the presence of multiple population roots for dimension $p = 2$, see Beran and Srivastava (1987). Their proof in the latter case makes use of the explicit form of the eigenvalues of a 2×2 matrix.

For the $p = 2$ dimensional case, Beran and Srivastava (1987) show that bootstrapping based upon samples of size m with $m/n \rightarrow 0$, gives a strongly consistent estimate of the limiting distribution of $n^{1/2}\{\varphi(S_n) - \varphi(\Sigma)\}$ irregardless of the eigenvalue multiplicities. This approach works in general. That is, suppose $\{x_i^*;\ 1 \leq i \leq m\}$ represent a random sample of size m from the distribution F_n, with $m/n \rightarrow 0$. Let $S_{(m)}^*$ be the sample covariance matrix of this sample, and let $X_{(m)}^* = m^{1/2}\{\varphi(S_{(m)}^*) - \varphi(S_n)\}$. The bootstrap estimate of the distribution is still strongly consistent. That is,

$$W_{(m)}^* = m^{1/2}(S_{(m)}^* - S_n) \rightarrow_{d^*} W \quad \text{a.s.} \tag{5.8}$$

Likewise, an analogous statement to (5.5) holds,

$$X_{(m)}^* - \{H(\tilde{W}_{(m)}^{*0} + \tilde{A}_{(m)}) - H(\tilde{A}_{(m)})\} \rightarrow_{d^*} 0 \quad \text{a.s.,} \tag{5.9}$$

where $W^{*0}_{(m)} = P_n W^*_{(m)} P'_n$ and hence $\tilde{W}^{*0}_{(m)}$ is defined accordingly. Also, $A_{(m)} = m^{1/2}(\Delta_n - \Delta)$ with $\tilde{A}_{(m)}$ defined accordingly. Now, $\tilde{W}^{*0}_{(m)} \to_{d^*} \tilde{W}$ a.s., and $A_{(m)} = (m/n)^{1/2} A_n \to_{a.s.} 0$, which by (5.8) gives

$$X_{(m)} = m^{1/2}\{\varphi(S^*_{(m)}) - \varphi(S_n)\} \to_{d^*} H(\tilde{W}) \quad \text{a.s.} \tag{5.10}$$

Although bootstrapping a sample of size $o(n)$ gives consistent results, its asymptotic efficiency is zero with respect to bootstrapping a sample of size n when the roots are distinct. A consistent and efficient method of bootstrapping eigenvalues which does not presuppose knowledge of the population eigenvalues multiplicity is an open problem.

It should be noted that the results of this section depend on the sample covariance matrix only through properties (5.1), (5.3) and (5.8). The results generalize to any symmetric estimate of Σ for which (5.1), (5.3) and (5.8) hold.

REFERENCES

Beran, R. and Srivastava, M. S. (1985). Bootstrap tests and confidence regions for functions of a covariance matrix. *Ann. Statist.* **13**, 95–115.

Beran, R. and Srivastava, M. S. (1987). Correction: Bootstrap tests and confidence regions for functions of a covariance matrix. *Ann. Statist.* **15**, 470–471.

Kato, T. (1980). *Perturbation Theory for Linear Operators.* Springer, Berlin.

Rao, C. R. (1973). *Linear Statistical Inference and its Applications.* Wiley, New York.

Tyler, D. E. (1983). The asymptotic distribution of principal component roots under local alternatives to multiple roots. *Ann. Statist.* **11**, 1232–1242.

Wielandt, H. (1967). Topics in the analytic theory of matrices. (Lecture notes prepared by R. R. Meyer.) University of Wisconsin Press, Madison.

p/t3